

Вычисление характеристик стационарных случайных последовательностей экстремальных величин

В статье рассмотрены методы оценивания характеристик стационарных случайных последовательностей экстремальных величин. В качестве математических моделей последовательностей экстремальных величин предложено использовать эконометрические модели $AR(1)$, $GARCH(1,1)$. В результате вычислительных экспериментов по сравнительному анализу классических эконометрических моделей с использованием нормального закона и обобщенного закона Парето показана эффективность предложенных автором эконометрических моделей для моделирования и оценивания характеристик стационарных случайных последовательностей экстремальных величин. Полученные результаты были использованы для моделирования волатильности как российского, так и зарубежных финансовых индексов.

1. Модель стационарных случайных последовательностей экстремальных величин

Рассмотрим стационарную случайную последовательность $(X_n, n \in N^+)$, которая задана на некотором стохастическом пространстве $(\Omega, \mathfrak{F}(\mathfrak{F}_n)_{n \geq 0}, P)$. Предположим, что динамика поведения X_n описывается следующим выражением:

$$X_n = \mu_n + \sigma_n Z_n, \quad (1)$$

причем в (1) $Z_n \approx F_z[0,1]$ является белым шумом в широком смысле. Мы предполагаем также, что условное среднее μ_n и волатильность σ_n \mathfrak{F}_{n-1} -измеримы. Найдем квантили таких распределений для $k = 1, \dots, N$, где N может быть достаточно большим. Для этого введем $F_X(x)$, функцию распределения процесса (X_n) , и пусть $F_{X_{n+1} + \dots + X_{n+k} | \mathfrak{F}_n}(x)$ условная функция распределения (X_n) за последующие k измерения на фильтрации \mathfrak{F}_n . Выражение для квантили условной функции распределения $F_{X_{n+1} + \dots + X_{n+k} | \mathfrak{F}_n}(x)$ имеет следующий вид: $x_q^n(k) = \inf\{x \in R: F_{X_{n+1} + \dots + X_{n+k} | \mathfrak{F}_n}(x) \geq q\}$.

Рассмотрим также условное математическое ожидание:

$$S_q^n(k) = E\left(\sum_{j=1}^k X_{n+j} \mid \sum_{j=1}^k X_{n+j} > x_q^n(k), \mathfrak{F}_n\right).$$

Поскольку

$$F_{X_{n+1} | \mathfrak{F}_n}(x) = P(\sigma_{n+1} Z_{n+1} + \mu_{n+1} \leq x | \mathfrak{F}_n) = F_Z((x - \mu_{n+1}) / \sigma_{n+1}),$$

то выражения для x_q^n и S_q^n имеют следующий вид:

$$x_q^n = \mu_{n+1} + \sigma_{n+1} Z_q, \quad (2)$$

$$S_q^n = \mu_{n+1} + \sigma_{n+1} E(Z | Z > z_q), \quad (3)$$

где z_q — квантиль функции распределения $F_Z(z)$. Мы должны построить соответствующие модели динамики для μ_n и σ_n прежде чем перейти к вычислению (2)–(3). Предложим модель AR(1) [Engle (1982)] для условного среднего μ_n в качестве таких моделей:

$$\mu_n = \theta X_{n-1} \quad (4)$$

и модель GARCH(1,1) [Bollerslev(1986)] для волатильности σ_n :

$$\sigma_n^2 = \alpha_0 + \alpha_1 \varepsilon_{n-1}^2 + \alpha_2 \sigma_{n-1}^2, \quad (5)$$

где $\varepsilon_n = X_n - \mu_n$, $\alpha_0 > 0$, $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_2 + \alpha_1 < 1$, $|\theta| < 1$.

Оценки условных средних $(\tilde{\mu}_1, \dots, \tilde{\mu}_n)$ и волатильностей $(\tilde{\sigma}_1, \dots, \tilde{\sigma}_n)$ могут быть вычислены последовательной подстановкой в выражения (4)–(5) при соответствующем выборе начальных значений.

Тогда значения остатков имеют следующий вид:

$$(z_1, \dots, z_n) = \left(\frac{x_1 - \tilde{\mu}_1}{\tilde{\sigma}_1}, \dots, \frac{x_n - \tilde{\mu}_n}{\tilde{\sigma}_n} \right).$$

Оценки параметров $\tilde{\theta} = (\tilde{\theta}, \tilde{\alpha}_0, \tilde{\alpha}_1, \tilde{\alpha}_2)$ могут быть получены методом максимального правдоподобия. Оценки для условного среднего $\tilde{\mu}_{n+1}$ и волатильности $\tilde{\sigma}_{n+1}$:

$$\tilde{\mu}_{n+1} = \tilde{\theta} x_n, \quad \tilde{\sigma}_{n+1}^2 = \tilde{\alpha}_0 + \tilde{\alpha}_1 \varepsilon_n^2 + \tilde{\alpha}_2 \tilde{\sigma}_n^2,$$

где $\tilde{\varepsilon}_n = x_n - \tilde{\mu}_n$.

Как показал анализ остатков для большинства финансовых временных серий [Pagan(1996)], функция распределения Z_n значительно отличается от предполагаемого в стандартных GARCH-моделях нормального распределения. На рис. 1 приведен график квантилей распределения остатков для дневных приращений значений фондового индекса S&P500 за период с 01.01.1996 по 31.10.2006 гг.

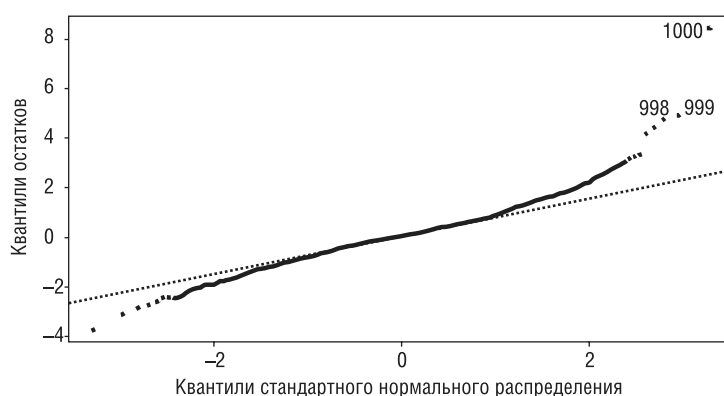


Рис. 1. График квантилей остатков относительно квантилей стандартного нормального распределения

Предложим в качестве модели, описывающей функцию распределения остатков z_k , обобщенное распределение Парето (GPD):

$$G_{\xi, \beta}(y) = \begin{cases} 1 - (1 + \xi y / \beta)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-y/\beta), & \xi = 0, \end{cases}$$

где $\beta > 0$, $y \geq 0$, когда $\xi \geq 0$, и $0 \leq y \leq -\beta/\xi$, $\xi < 0$.

В работах Балкемы и Хаана [Balkema, Naan (1974)], [Pickands (1975)] доказано, что функции распределения $F_Z(z)$ из рассматриваемого нами класса аппроксимируются обобщенным распределением Парето с параметром $\xi > 0$ высокой степени точности. Положим, что $N = \sum_{i=1}^k 1\{z_i > u\}$ — случайное число, равное количеству значений z_k из выборки остатков, превысивших порог u . Тогда имеет место соотношение

$$1 - F_Z(z) = (1 - F(u))(1 - F_u(z - u)). \quad (6)$$

Из (6) получаем выражение для функции $\tilde{F}_Z(z) = P(Z > u)$:

$$\tilde{F}_Z(z) = 1 - \frac{N}{n} (1 + \tilde{\xi}_N(z - u) / \tilde{\beta}_N)^{-1/\tilde{\xi}_N}, \quad (7)$$

где $\tilde{\xi}_N, \tilde{\beta}_N$ — оценки значений параметров GPD, полученные методом максимального правдоподобия. Пусть нами построена порядковая статистика остатков z_1, z_2, \dots, z_n . Зафиксируем некоторое число $N = k, k \ll n$ и таким образом получим случайный порог $u = z_{k+1}$. Построим выражение (7) для выборки остатков следующего вида $(z_1 - z_{k+1}, \dots, z_k - z_{k+1})$:

$$\tilde{F}_Z(z) = 1 - \frac{k}{n} \left(1 + \tilde{\xi}_k \frac{z - z_{k+1}}{\tilde{\beta}_k} \right)^{-1/\tilde{\xi}_k}. \quad (8)$$

Для $q > 1 - k/n$ получим выражение для \tilde{z}_q :

$$\tilde{z}_q = z_{k+1} + \frac{\tilde{\omega}_k}{\tilde{\xi}_k} \left(\left(\frac{1 - q}{k/n} \right)^{-\tilde{\xi}_k} - 1 \right). \quad (9)$$

Учитывая (7), (8), (9) для используемых нами данных фондового индекса S&P500, мы построили оценки при значении $k = 100$. Результаты показаны на рис. 2. На нем для сравнения приведены аналогичные оценки для нормального распределения и t_4 -распределения Стьюдента.

Из сравнительного анализа полученных результатов следует, что нормальное распределение в значительной степени недооценивает как эксцессивные убытки, так и эксцессивные прибыли для исследуемых данных. Распределение GPD дает достаточно хорошие аппроксимации как для положительных, так и для отрицательных значений приращений остатков [Щетинин, Лапушкин (2004)].

2. Вычисление оценок моментов функции распределения остатков

Для построения оценок (2)–(3), мы должны вычислить $E(Z | Z > z_q)$ для функции распределения $F_Z(z)$. Положим, что

$$Z - u | Z > u \sim GPD | (\xi, \beta),$$

где $\beta + u\xi > 0$, $\xi < 1$. Получаем

$$E(Z | Z > u) = \frac{u + \beta}{1 - \xi}. \quad (10)$$

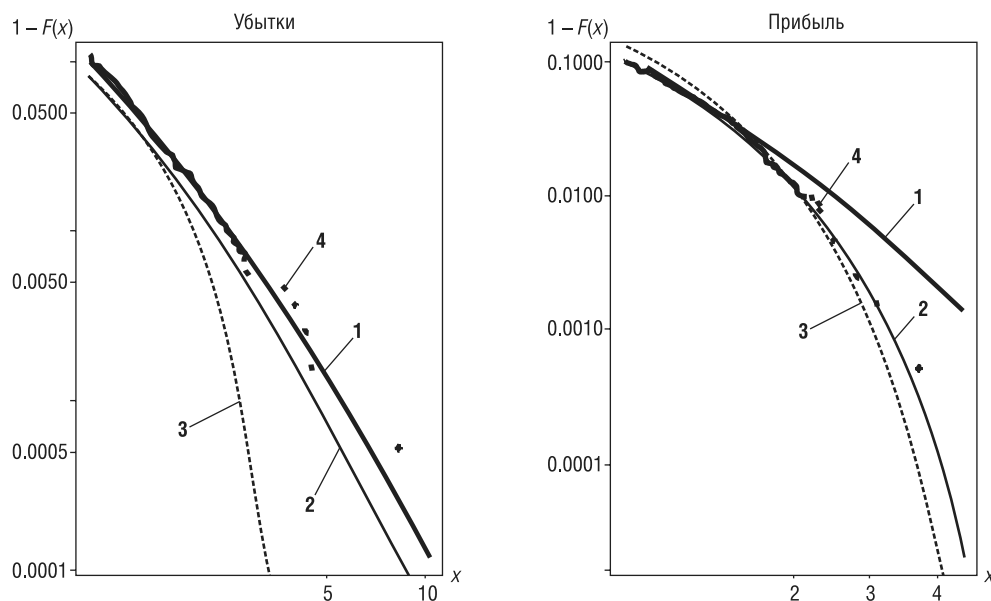


Рис. 2. Аппроксимация функции распределения дневных логарифмических приращений индекса S&P500 за период с 01.01.1996–31.10.2006 гг.

график 1 — обобщенное распределение Парето (GPD), график 2 — t_4 -распределение Стюдента, график 3 — стандартное нормальное распределение, график 4 — эмпирическое распределение. (Все оси представлены в логарифмическом масштабе.)

Для $z_q > u$ можно записать:

$$Z - z_q | Z > z_q \approx GPD(\xi, \beta + \xi(z_q - u)). \quad (11)$$

Отсюда, воспользовавшись (10), получим

$$E(Z | Z > z_q) = z_q \left(\frac{1}{1 - \xi} + \frac{\beta - \xi u}{(1 - \xi)z_q} \right). \quad (12)$$

Воспользовавшись выражением (9) и (12) для оценки \tilde{z}_q , а также, приняв $u = z_{k+1}$, получим оценку \tilde{S}_q^n :

$$\tilde{S}_q^n = \tilde{\mu}_{n+1} + \tilde{\sigma}_{n+1} \left(\frac{1}{1 - \tilde{\xi}_k} + \frac{\tilde{\beta}_k - \tilde{\xi}_k \tilde{z}_{k+1}}{(1 - \tilde{\xi}_k) \tilde{z}_k} \right). \quad (13)$$

Исходя из (13), построим оценку отношения S_q^n и x_q^n :

$$\frac{S_q^n}{x_q^n} \approx \frac{S_q^n - \mu_{n+1}}{x_q^n - \mu_{n+1}} = \frac{E(Z | Z > z_q)}{z_q}. \quad (14)$$

В случае, когда $F_Z(z)$ является стандартным нормальным распределением:

$$E(Z | Z > z_q) = k(z_q),$$

где $k(x) = \varphi(x) / (1 - \Phi(x))$ — отношение Милла, а $\varphi(x)$ и $\Phi(x)$ — плотность и функция стандартного нормального распределения соответственно. Можно воспользоваться в этом случае асимптотическим выражением для $k(x) = x(1 + x^{-2} + O(x^{-2}))$, при $x \rightarrow \infty$, из которого следует, что $\frac{S_q^n}{x_q^n} \rightarrow 1$ при $q \rightarrow 1$. Для GPD это отношение сводится к $(1 - \xi)^{-1} > 1$, поскольку $k(x) \approx x(1 + (\sqrt{1 + 8/x^2})/4)$. В табл. 1 приведены соответствующие значения для отношения (14).

Таблица 1

Значения отношения (14) для обобщенного распределения Парето $\varsigma \tilde{\xi} = 0,231$, $\tilde{\omega} = 0,564$, при $u = z_{k+1} = 1,212$, и нормального распределения остатков

Тип распределения	Порядок квантили q			
	0,95	0,99	0,995	$q \rightarrow 1$
Обобщенное распределение Парето (GPD)	1,506	1,418	1,372	1,26
Нормальное распределение (Ф)	1,25	1,15	1,11	1,00

Из табл. 1 видно, что использование нормального распределения для описания функции распределения $F_Z(z)$ приводит к занижению оценок \tilde{x}_q^n и \tilde{S}_q^n для $q \geq 0,95$. Нормированные остатки определим в следующем виде:

$$R_{n+1} = \frac{X_{n+1} - S_q^n}{\sigma_{n+1}} = Z_{n+1} - E(Z|Z > z_q), R_{n+1}^* = R_{n+1} / \sqrt{\text{var}(Z|Z > z_q)},$$

где $Z - u | Z > u \sim \text{GPD} | (\xi, \beta)$,

учитывая (11), имеем величину риска (VaR):

$$\text{var}(Z|Z > z_q) = \frac{(\beta + \xi(z_q - u))^2}{(1 - \xi)^2(1 - 2\xi)} \quad (15)$$

для $z_q > u$ и $\xi < 0,5$. Аналогичное выражение для (15) в случае нормального распределения Z с использованием отношения Милла имеет вид:

$$\text{var}(Z|Z > z_q) = 1 + z_q k(z_q) - k(z_q)^2. \quad (16)$$

Нас будут интересовать различия между значениями X_{n+1} и оценками \tilde{S}_q^n в случае, когда $X_{n+1} > \tilde{x}_q^n$, что эквивалентно $Z_{n+1} > \tilde{z}_q$. Сформируем выборку значений x_{n+1} таких, что $x_{n+1} > \tilde{x}_q^n$. Для них, учитывая (16), вычислим остатки $\{r_{n+1}^*: n \in N, x_{n+1} > \tilde{x}_q^n\}$, где

$$r_{n+1}^* = \frac{x_{n+1} - \tilde{S}_q^n}{\tilde{\sigma}_{n+1} \sqrt{\text{var}^*(Z|Z > z_q)}}, \quad (17)$$

а значения \tilde{S}_q^n и $\text{var}^*(Z|Z > z_q)$ вычисляются последовательно для каждого x_{n+1} по формулам (13) и (15) для GPD и по формуле (16) для нормального распределения. На основании проведенных вычислений r_n^* для GPD и нормального распределения допущение о нормальности экстремальных остатков r_n^* является, на наш взгляд, несостоятельным.

По имеющимся у нас оценкам \tilde{X}_q^n рассмотрим оценки условных квантилей $\tilde{X}_q^n(k)$ для $k > 1$. Получить выражение для них подобно (9), (13) — достаточно сложно, поскольку GARCH-модели для распределений $F_{X_{n+1}+\dots+X_{n+k}|\mathfrak{I}_n}(x)$ неизвестны в аналитическом виде [Щетинин, Назаренко (2004)]. Поэтому мы применили следующий подход: аппроксимировали выборку наших данных моделью AR(1)-GARCH(1,1) (4)–(5), а затем оценили «хвосты» распределения $\tilde{F}_Z(z)$ из выражения (8). Мы разработали процедуру моделирования функции распределения $\tilde{F}_Z(z)$ в следующем виде:

$$\tilde{F}_Z(z) = \begin{cases} \frac{k}{n} \left(1 + \xi_k^{(2)} \frac{|z - z_{n-k}|}{\beta_k^{(2)}} \right)^{-1/\xi_k^{(2)}}, & z < z_{n-k}, \\ \frac{1}{n} \sum_{i=1}^n 1(z_i \leq z), & z_{n-k} \leq z \leq z_{k+1}, \\ 1 - \frac{k}{n} \left(1 + \xi_k^{(1)} \frac{z - z_{k+1}}{\beta_k^{(1)}} \right)^{-1/\xi_k^{(1)}}, & z > z_{k+1}. \end{cases}$$

Можно построить L выборок $(x_{n+1}, \dots, x_{n+k})$ и, используя эту модель и оценки параметров модели (4) и (5), вычислить соответствующие кумулятивные суммы реализаций $\sum_{j=1}^k X_{n+j}|\mathfrak{I}_n$, откуда затем вычислить значения квантилей $\tilde{X}_q^n(k)$. Для значений $L \approx 10000$ мы получили достаточно хорошие оценки для $q \geq 0,95$.

С тем чтобы выявить эмпирический закон зависимости условной квантили $\tilde{X}_q^n(k)$ от \tilde{X}_q^n , рассмотрим $\tilde{X}_q^n(k)/\tilde{X}_q^n, k > 1$. Феллер [Феллер (1984)] установил, что для независимых, одинаково распределенных случайных величин X_i , имеющих «тяжелохвостый» тип распределения:

$$(x^{-1/\xi} k L(x))^{-1} P(X_1 + \dots + X_k > x) \rightarrow 1 \text{ при } x \rightarrow \infty.$$

Отсюда можно получить для квантилей высокого порядка $q \geq 0,95$

$$\tilde{X}_q^n(k)/\tilde{X}_q^n \approx k^{\lambda(n)}, \quad (18)$$

где $\lambda(n)$ зависит от значения σ_n . При $k \rightarrow \infty$ отношение (18) принимает вид $\tilde{X}_q^n(k)/\tilde{X}_q^n \approx k^{1/2}$ [Balkema, Naan (1974)]. Мы применили наш алгоритм [Стихова (2006)] вычисления квантилей $\tilde{X}_q^n(k)$ для $k = 1, \dots, 50$ и различных уровней волатильности σ_n . Табл. 2 содержит полученные значения $\lambda(n)$.

Таблица 2

Значения $\lambda(n)$ для $k = 1, \dots, 50$ и различных уровней волатильности σ_n

Уровни волатильности σ_n	Порядок квантили q	
	0,95	0,99
Низкий	0,643	0,646
Средний	0,60	0,581
Высокий	0,481	0,473

Анализ табл. 2 показывает, что если начальное значение σ_n велико (мало), то будущие значения волатильности, будут в среднем меньше (больше), чем σ_n . Таким образом, условная квантиль $\tilde{\chi}_q^n(k)$ будет возрастать медленнее (соответственно быстрее) по k . Рис. 3 иллюстрирует соответствующие результаты для $q = 0,99$.

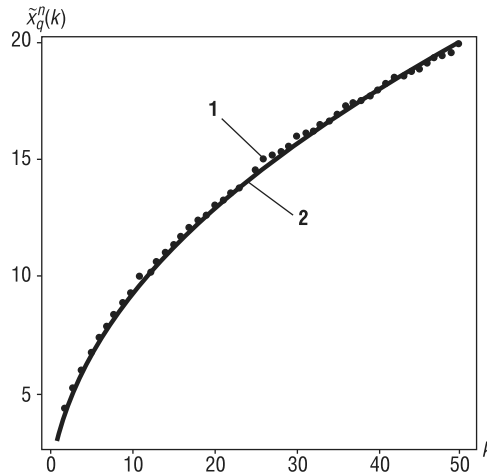


Рис. 3. Тестирование закона (18) для оценок эмпирических условных квантилей $\tilde{\chi}_q^n(k)$ в зависимости от k

график 1 — эмпирическое распределение, график 2 — соответствующие условные квантили, вычисленные по формуле (18)

В проведенных нами экспериментах по моделированию волатильности фондовых индексов NASDAQ, PTC, DAX-30 и ее прогнозированию, используя нашу модель (5) за периоды с 1973 по 2006 гг., мы получили результаты, представленные на рис. 4 и в таблицах 3–6. Согласно полученным результатам, наша модель достаточно точно совпадает с исходными данными.

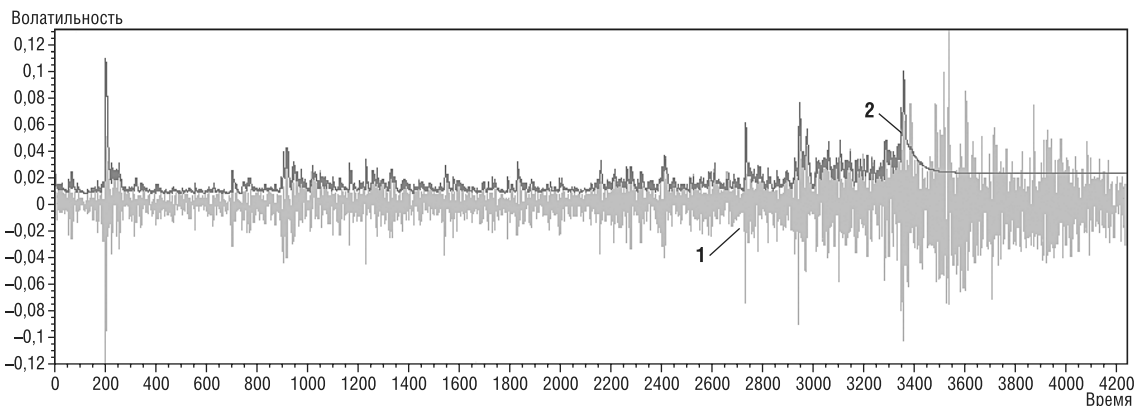


Рис. 4. Прогноз волатильности фондового индекса NASDAQ с использованием модели (5) за период 01.01.1973–31.10.2006 гг.

график 1 — оценка волатильности логарифмических приращений индекса NASDAQ, график 2 — модель (5) волатильности и ее прогноз

Таблица 3

Оценки квантилей $\tilde{\chi}_q$ для логарифмических приращений дневных цен закрытия фондового индекса РТС за период 10.01.1998–31.10.2006 гг. для различных значений порога u

	u	k	$\tilde{\xi}$	$q = 0,999$	$q = 0,9999$
$\tilde{\chi}_q$	0,05	729	0,0369	[0,198, 0,224, 0,261]	[0,32, 0,37, 0,426]
	0,06	597	0,0557	[0,206, 0,238, 0,321]	[0,349, 0,396, 0,47]
	0,08	361	0,0985	[0,227, 0,26, 0,348]	[0,35, 0,43, 0,485]
	0,1	239	0,1168	[0,23, 0,26, 0,36]	[0,373, 0,43, 0,55]

Таблица 4

Оценки квантилей $\tilde{\chi}_q$ для логарифмических приращений дневных цен закрытия фондового индекса NASDAQ за период 01.01.1973–31.10.2006 гг. для различных значений порога u

	u	k	$\tilde{\xi}$	$q = 0,99999$	$q = 0,999999$
$\tilde{\chi}_q$	0,01	834	0,217	[0,101, 0,122, 0,143]	[0,187, 0,241, 0,29]
	0,015	414	0,158	[0,101, 0,122, 0,143]	[0,19, 0,249, 0,307]
	0,02	227	0,118	[0,102, 0,126, 0,15]	[0,195, 0,25, 0,316]

Таблица 5

Оценки квантилей $\tilde{\chi}_q$ для логарифмических приращений дневных цен закрытия фондового индекса DAX-30 за период 01.01.1993–12.05.2006 гг. при различных значениях порога u

	u	k	$\tilde{\xi}$	$q = 0,999$	$q = 0,9999$
$\tilde{\chi}_q$	0,005	914	0,00062	[0,038, 0,063, 0,089]	[0,032, 0,09, 0,149]
	0,01	546	0,03938	[0,033, 0,064, 0,096]	[0,024, 0,093, 0,161]
	0,02	163	0,0116	[0,021, 0,069, 0,117]	[0,005, 0,094, 0,194]
	0,03	48	0,0077	[0,044, 0,069, 0,094]	[0,058, 0,096, 0,127]

Таблица 6

Оценки значений и 95%-ые доверительные интервалы для величины риска q ($Var(q)$) и ожидаемого дефицита q ($ES(q)$) для отрицательных логарифмических приращений дневных цен закрытия фондового индекса РТС за период 10.01.1998–31.10.2006 гг. при значении порога $u = 0,06$

q	Величина риска q ($Var(q)$)		Ожидаемый дефицит q ($ES(q)$)	
	Значение	Доверительный интервал	Значение	Доверительный интервал
0,99	0,111	[0,103, 0,123]	0,17	[0,169, 0,172]
0,995	0,152	[0,129, 0,215]	0,21	[0,197, 0,221]
0,999	0,238	[0,206, 0,321]	0,322	[0,232, 0,387]
0,9999	0,32	[0,314, 0,382]	0,435	[0,337, 0,539]

3. Заключение

В работе рассмотрены методы вычисления условных квантилей и математических ожиданий стационарных случайных последовательностей экстремальных величин. Для моделирования функции условного распределения инновационного процесса Z_n мы использовали обобщенное распределение Парето. Для описания динамики характеристик последовательностей мы использовали параметрические модели AR(1) и GARCH(1,1). В вычислительных экспериментах исследовались индексы S&P500, PTC, NASDAQ за период с 1973 по 2006 гг., и было показано, что предложенный подход является более точным для моделирования волатильности, чем общепринятый метод моделирования функции распределения процесса Z_n с помощью стандартного нормального распределения.

Предложена модификация эмпирического закона скейлинга (18) условных квантилей $x_q^n(k)$ для $k > 1$ последующих надпороговых значений $(X_{n+1}, \dots, X_{n+k})$. Приведенные в работе вычислительные эксперименты показали соответствие отношения (18) исследованным нами финансовым временным сериям.

Результаты могут быть использованы для моделирования волатильности финансовых активов и оценивания рисков инвестирования.

Список литературы

Стихова О. В. Моделирование, оценка и методы прогнозирования рыночных рисков / XLII Всероссийская конференция по проблемам математики, информатики, физики и химии. Тезисы докладов. Секции физики. М.: Изд-во РУДН, 2006. С. 77.

Стихова О. В. Оценка рыночных рисков. Финансовый анализ. Методы прогнозирования / XIII конференция серии Математика. Компьютер. Образование / Сборник научных тезисов под ред. Г. Ю. Ризниченко, НИЦ «Регулярная и хаотическая динамика». Ижевск, 2006. С. 286–287.

Феллер В. Введение в теорию вероятностей и ее приложения. М: Мир, 1984.

Щетинин Е. Ю., Лапушкин А. С. Статистические методы и математические модели оценивания финансовых рисков. Математическое моделирование. Т. 16. Вып. 5. М.: Наука, 2004.

Щетинин Е. Ю., Назаренко К. М., Парамонов А. В. Инструментальные методы стохастического анализа экстремальных событий. Математическое моделирование и оптимальное управление // Вестник ННГУ, Н. Новгород, 2004. №2(29).

Balkema A., Haan L. de. Residual life time at great age. Annals of Probability. 1974. № 2.

Bollerslev T. Generalised Autoregressive Conditional Heteroscedastity // J. of Econometrics. 1986. № 31.

Engle R. Autoregressive conditional heteroscedastity with estimates of the variance of United Kingdom inflation. Econometrica. 1982. № 4.

Pagan A. The econometrics of financial markets // J. of Empirical Finance. 1996. № 3.

Pickands J. Statistical inference using extreme order statistics // The Annals of Statistics. 1975. № 3.